



## Profilométrie par déphasage en mouvement pour applications industrielles

Pierre Schroeder, Romain Roux, Jean-Marie Favreau, Mathieu Perriollat, Adrien Bartoli

### ► To cite this version:

Pierre Schroeder, Romain Roux, Jean-Marie Favreau, Mathieu Perriollat, Adrien Bartoli. Profilométrie par déphasage en mouvement pour applications industrielles. Orasis, Congrès des jeunes chercheurs en vision par ordinateur, Jun 2013, Cluny, France. hal-00829432

**HAL Id: hal-00829432**

**<https://hal.science/hal-00829432>**

Submitted on 5 Jun 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Profilométrie par déphasage en mouvement pour applications industrielles

Pierre Schroeder<sup>1</sup> Romain Roux Jean-Marie Favreau<sup>1</sup> Mathieu Perriollat Adrien Bartoli<sup>1</sup>

<sup>1</sup> ISIT – Image Science for Interventional Techniques, UMR 6284 Uda – CNRS

28, place Henri Dunant, 63001 Clermont-Ferrand, France  
pierre.schroeder@etu.udamail.fr

## Résumé

*La profilométrie par déphasage est une méthode éprouvée pour reconstruire des surfaces de manière dense et précise. Par contre la scène doit rester immobile pendant l'acquisition d'une séquence de plusieurs images. De plus, il existe des méthodes de stéréo active qui s'affranchissent de la contrainte d'immobilité de la scène mais qui imposent d'autres limitations comme par exemple la continuité de la surface et de la texture ou une résolution de reconstruction considérablement réduite.*

*Nous présentons une nouvelle technique de reconstruction aussi dense et précise que la profilométrie par déphasage et qui permet une translation de la scène pendant l'acquisition de la séquence d'images. Cela la rend intéressante pour des applications industrielles. Nous étudions sa performance à l'aide de simulations et donnons une démonstration sur un exemple réel.*

## Mots Clef

Reconstruction 3D, stéréo active, déphasage, profilométrie, métrologie, inspection industrielle, projection de franges

## Abstract

*Phase-Shift Profilometry (PSP) provides a means for dense high-quality surface scanning. However it imposes a staticity constraint : The scene is required to remain still during the acquisition of multiple images. On the other hand, there exist active stereo techniques which overcome these constraints but impose other limitations, for instance on the surface's continuity or texture, or by significantly reducing the reconstruction's resolution.*

*We present a novel approach to recover reconstructions as dense and almost as accurate as PSP, but which allows for a translational object/scene motion during the acquisition of multiple input frames ; and study its performance in simulations and present real data results.*

## Keywords

3D reconstruction, active stereo, phase-shift, profilometry, metrology, industrial inspection, fringe projection

## 1 Introduction

Les méthodes de reconstruction stéréo actives représentent un moyen d'acquisition 3D de surfaces peu ou pas texturées n'utilisant qu'un nombre limité d'hypothèses concernant la surface à reconstruire. Les approches alternatives à la stéréo active, sont la stéréo passive (multi-caméras) [9] qui requièrent des surfaces suffisamment texturées et les caméras temps de vol [5] qui fournissent des reconstructions plus bruitées et à résolution réduite.

On peut également distinguer de nombreux types de stéréo active : par exemple, la surveillance respiratoire de patients requiert un système temps-réel mais pas nécessairement une grande robustesse vis-à-vis de la texture, ni une densité de reconstruction importante. Les applications multimédia (e.g. industrie du jeu vidéo) ciblent l'usage exclusivement visuel des modèles 3D reconstruits et se contentent donc de reconstructions éparées dont la texture compense le manque de détails géométriques, par contre l'acquisition doit pouvoir se faire avec un scanner 3D portable. L'archivage numérique d'ouvrages historiques (e.g. bustes et reliefs) demande en revanche une reconstruction dense de haute qualité, mais satisfait généralement la contrainte d'immobilité de la scène.

Dans le domaine de la métrologie et de l'inspection industrielle les méthodes s'appuient souvent sur les approches de Profilométrie par Déphasage (PDP) d'un motif projeté [10]. Celles-ci génèrent des reconstructions denses et précises indépendamment de la présence ou absence de texture dans la scène. Par contre, celles-ci imposent que la scène reste immobile pendant l'acquisition de plusieurs images consécutives en déphasant le motif projeté ou requièrent un système entièrement équipé d'optiques télécentriques (caméras et projecteurs) très coûteuses qui supporte nativement d'utiliser le mouvement. Les techniques qui n'imposent pas cette contrainte d'immobilité n'atteignent que des résolutions de reconstruction réduites car elles sont limitées par la résolution native du vidéo-projecteur et donnent naissance à des problématiques de fusion de multiples reconstructions éparées [3], rendent le système sensible à la texture [4] ou à la couleur [14] de la scène, ou demandent du matériel haute-fréquence spécifique [12]. Beaucoup de systèmes d'inspection industriels requièrent

la densité fournie par la PDP mais ne peuvent pas respecter la contrainte d'immobilité de la scène. Ceci est par exemple le cas pour mesurer des objets convoyés sur une chaîne de fabrication. En revanche les mouvements que peuvent subir les objets entre deux images dans ce contexte sont connus et limités. Ce cas particulier n'étant pas couvert dans la littérature, nous avons décidé de développer la méthode présentée dans cet article. Notre méthode peut reconstruire de façon dense une scène en 3D avec une précision proche de celle atteinte par la PDP. Elle reste robuste à la texture et la couleur de la scène et ne requiert pas du matériel haute-fréquence.

**Plan.** La Partie 2 est une synthèse de l'état de l'art. La Partie 3 présente notre approche. Ses performances sont étudiées à travers une scène de synthèse dans la Partie 4 et une scène réelle dans la Partie 5. Les résultats sont commentés dans la Partie 6 et nos conclusions présentées dans la Partie 7.

## 2 État de l'art

**Méthodes à motifs binaires.** Les méthodes à projection unique de motifs binaires essayent de réduire les ambiguïtés de décodage [3] ou de maximiser la discrimination de disparité du motif projeté (pour des systèmes multi-caméra) [6]. Si les scènes ne contiennent que peu de texture et des couleurs neutres, les systèmes avec projecteur couleur et caméra couleur permettent de multiplexer plusieurs signaux déphasés ou binaires sur les canaux RGB [14].

**Méthodes à motifs continus.** Weise, Leibe et Van Gool [12] fixent instantanément la scène en utilisant un système de profilométrie à trois déphasages à haute fréquence. Les vitesses et accélérations estimées des objets grossièrement reconstruits sont ensuite mises en relation avec l'erreur d'estimation de phase. Les reconstructions initiales sont ensuite corrigées grâce à une table de correspondance précalculée de manière analytique qui permet de corriger l'erreur d'estimation de phase. Les approches spatio-temporelles [13, 1] cherchent à valider dans le temps les observations spatiales.

Une vision globale plus exhaustive des méthodes de stéréo active est donnée par Salvi et al. [7].

Même si les méthodes susmentionnées permettent d'obtenir de bons résultats, elles ne ciblent pas des applications industrielles. Les techniques qui s'appuient sur des motifs binaires sont limitées par les caractéristiques du projecteur (résolution native) et requièrent des fenêtres de décodage ou de mise en correspondance de taille importante (plusieurs pixels projecteur), ce qui réduit d'autant plus le niveau de détails atteignable. Les méthodes d'estimation de phase à image unique ont des difficultés aux discontinuités de géométrie et de texture.

Notre approche répond à une demande industrielle (non explicitée pour des raisons de confidentialité) dans laquelle la contrainte d'immobilité est trop coûteuse à réaliser.

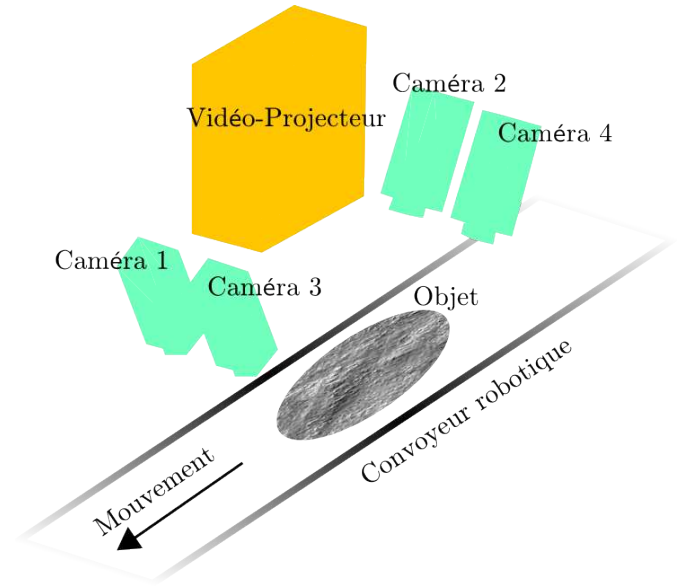


FIG. 1: Représentation schématique de notre mise en œuvre matérielle.

## 3 Notre Approche

### 3.1 Mise en œuvre matérielle

Notre système d'acquisition, illustré par le schéma de la Figure 1, est composé d'une partie optique et d'une partie mécanique. Il est isolé des sources de lumières externes et ne contient aucun éclairage ambiant. La partie optique contient un vidéo-projecteur SVGA ou XGA éclairant la scène ainsi que quatre caméras synchronisées de résolution de 4 millions de pixels et équipées d'optiques classiques (i.e. non-télécentriques). Le système peut aussi être réalisé avec deux caméras au lieu de quatre. Les caméras sont calibrées [15] et positionnées de sorte que leurs Champs de Vue (CdV) se superposent au mieux dans le Champ de Profondeur (CdP) ciblé.

La partie mécanique est un convoyeur de précision robotique imposé par d'autres étapes de fabrication. Il fournit une mesure précise des mouvements.

### 3.2 Choix du motif projeté

Afin de bien choisir entre la projection d'un motif binaire ou continu, il faut prendre en compte que les caméras ont une résolution considérablement plus fine que les vidéo-projecteurs, l'utilisation de motifs binaires réduit considérablement la résolution de reconstruction car le projecteur est le facteur limitant. Cela est d'autant plus vrai que le décodage ou la mise en correspondance requièrent un voisinage minimum de plusieurs pixels du motif.

Les motifs continus en revanche, et les franges sinusoïdales en particulier, ont été utilisés avec succès dans la littérature existante pour réaliser de manière simple la mise en correspondance sub-pixelique et en maximisant l'information le long des lignes épipolaires. Nous avons donc choisi de projeter des franges sinusoïdales.

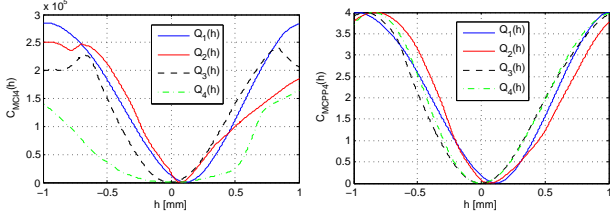


FIG. 2: Fonctions de coût MCI4 (gauche) et MCPPr (droite) de quatre points sur les images de synthèse (sans bruit caméra ni d'incertitude de mouvement).

### 3.3 Référence des points en mouvement

Comme les objets à reconstruire représentent des surfaces relativement plates, nous définissons l'échantillonnage de reconstruction comme une grille de points coplanaires. Cela donne  $m$  points  $\mathbf{Q}_i^1(h_i) = [x_i \ y_i \ h_i]^\top$  dans l'espace en fonction de leur hauteur  $h_i$ . Avec  $\mathbf{R}_t$  et  $\mathbf{t}_t$  représentant les rotations et translations tri-dimensionnelles qui réalignent les points au moment  $t = 1$  à eux-mêmes à l'instant  $t$ , les points correspondants  $\mathbf{Q}_i^t$  à l'instant  $t$  suivent la relation :

$$\mathbf{Q}_i^t(h_i) = \mathbf{R}_t \mathbf{Q}_i^1(h_i) + \mathbf{t}_t. \quad (1)$$

Les coordonnées d'un pixel d'un point dans le temps sont calculées pour chaque caméra  $C$  avec sa matrice de projection  $\mathbf{P}_C$  correspondante :

$${}^C \mathbf{q}_i^t(h_i) \sim \mathbf{P}_C \mathbf{Q}_i^t(h_i). \quad (2)$$

Le point placé sous une lettre fait référence aux coordonnées homogènes et  $\sim$  représente l'égalité à un facteur près. Le projecteur et les caméras sont calibrés et la distortion radiale est corrigée dans les images, ce qui permet de décrire la géométrie de la construction précisément à l'aide du modèle sténopé.

En appelant  $\mathcal{I}_t^C(\mathbf{q})$  l'intensité interpolée de manière bilinéaire au pixel  $\mathbf{q}$  de la  $t^{\text{ème}}$  image parmi  $n$ , acquise par la caméra  $C$ , nous construisons le vecteur :

$$\mathcal{I}_i^C(h_i) = \begin{bmatrix} \mathcal{I}_1^C({}^C \mathbf{q}_i^1(h_i)) & \dots & \mathcal{I}_n^C({}^C \mathbf{q}_i^n(h_i)) \end{bmatrix}^\top. \quad (3)$$

Si la hauteur  $h_i$  du point de la surface était connue, ce vecteur contiendrait les niveaux de gris du même point de l'objet. Pour toute autre hauteur, le vecteur présente une erreur de parallaxe. En outre, l'illumination d'un point physique varie entre les acquisitions consécutives, à cause de la variation spatiale du motif projeté et du déplacement de la scène par rapport à la projection. Par conséquent, le même point est imagé avec des intensités différentes pendant toute une séquence d'images.

Ensuite, nous utilisons les fonctions vectorielles  $\mathcal{I}_i^C(h_i)$  pour définir des fonctions de coût  $\mathcal{C}$ , que nous développons dans les Sections 3.4 et 3.5. Ces coûts font l'objet d'une minimisation non-linéaire en fonction des hauteurs des points de la grille de référence. Pour  $r$  caméras nous écrivons le

problème de reconstruction 3D comme :

$$\arg \min_{\hat{h}_i} \mathcal{C}(\hat{h}_i) = \arg \min_{\hat{h}_i} \mathcal{C}'(\mathcal{I}_i^1(\hat{h}_i), \dots, \mathcal{I}_i^r(\hat{h}_i)) \quad (4)$$

pour chaque point indépendamment.

### 3.4 Mise en correspondance par intensités

La Mise en Correspondance par Intensités (MCIr) n'utilise aucun *a priori* concernant le motif projeté et se base directement sur la différence d'intensité pour mettre en correspondance un point dans plusieurs caméras sur toute une séquence d'images. Pour  $r \geq 2$  caméras le coût s'écrit :

$$\mathcal{C}_{\text{MCIr}}(h_i) = \sum_{k=1}^r \left\| \mathcal{I}_i^k(h_i) - \frac{1}{r} \sum_{l=1}^r \mathcal{I}_i^l(h_i) \right\|_2^2. \quad (5)$$

### 3.5 Mise en correspondance par pseudo-phase

La Mise en Correspondance par Pseudo-Phase (MCPPr) est construite par analogie au principe de la PDP. La différence importante est que le déphasage est exercé par le déplacement de la scène plutôt que par le déphasage du motif projeté. Le motif projeté est défini pour un pixel  $\mathbf{p}$  par :

$$\mathcal{L}(\mathbf{p}) \sim \frac{1}{2} + \frac{1}{2} \cos(\boldsymbol{\omega}^\top \mathbf{p}) \quad (6)$$

avec le vecteur  $\boldsymbol{\omega}$  de taille deux, qui définit la fréquence des franges projetées.

Nous supposons que, pour un point physique de la surface, la relation entre l'éclairage par le projecteur et l'intensité acquise d'une caméra, telle qu'elle est modélisée dans la PDP, reste valide dans notre cas, ce que nous écrivons sous la forme :

$$\mathcal{I}_t^C({}^C \mathbf{q}_i^t(h_i)) = A_i + B_i \cos(\boldsymbol{\omega}^\top \mathbf{p}_i^t(h_i)) \quad (7)$$

avec le pixel projecteur  $\mathbf{p}_i^t(h_i) \sim \mathbf{P}_P \mathbf{Q}_i^t(h_i)$ . Ce qui peut être réécrit sous la forme :

$$\mathcal{I}_t^C({}^C \mathbf{q}_i^t(h_i)) = A_i + B_i \underbrace{\cos(\boldsymbol{\omega}^\top \mathbf{p}_i^1(h_i))}_{\varphi_i(h_i)} + \underbrace{\boldsymbol{\omega}^\top (\mathbf{p}_i^t - \mathbf{p}_i^1(h_i))}_{\delta \varphi_i^t(h_i)}. \quad (8)$$

Les équations (7) et (8) introduisent la notion de déphasage par changement de point de vue, par opposition à un déphasage par projection [10]. Elles conservent pour un point les propriétés passives : la texture/couleur  $A$  et un coefficient  $B$  modélisant l'amplitude de l'illumination modulée par le motif sinusoïdal. En notation matricielle on obtient :

$$\mathcal{I}_i^C(h_i) = \Delta_i(h_i) \mathbf{x}_i^C(h_i), \quad (9)$$

avec :

$$\Delta_i(h_i) = \begin{bmatrix} 1 & \cos(\delta \varphi_i^1(h_i)) & -\sin(\delta \varphi_i^1(h_i)) \\ \vdots & \vdots & \vdots \\ 1 & \cos(\delta \varphi_i^n(h_i)) & -\sin(\delta \varphi_i^n(h_i)) \end{bmatrix} \quad (10)$$

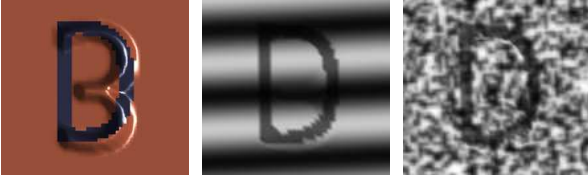


FIG. 3: Visualisation de la scène virtuelle : un plan avec un "3" en relief et un "D" en couleur (gauche), région d'intérêt d'une image synthétique avec projection de franges sinusoidales (milieu) et avec projection d'un motif binaire aléatoire.

et :

$$\mathbf{x}_i^C(h_i) = \begin{bmatrix} A \\ B \cos(\varphi_i(h_i)) \\ B \sin(\varphi_i(h_i)) \end{bmatrix}. \quad (11)$$

L'estimateur  $\hat{\mathbf{x}}_i^C(h_i) = [\hat{\gamma}_i^C(h_i) \ \hat{\alpha}_i^C(h_i) \ \hat{\beta}_i^C(h_i)]^T$  pour  $\mathbf{x}_i^C(h_i)$  peut être calculé à l'aide de la pseudo-inverse de Moore-Penrose  $\cdot^\dagger$  par :

$$\hat{\mathbf{x}}_i^C(h_i) = \Delta_i^\dagger(h_i) \mathcal{I}_i^C(h_i), \quad (12)$$

et dont les composants servent à déterminer :

$$\begin{bmatrix} \hat{c}_i^C(h_i) \\ \hat{s}_i^C(h_i) \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_i^C(h_i) \\ \hat{\beta}_i^C(h_i) \end{bmatrix} / \left\| \begin{bmatrix} \hat{\alpha}_i^C(h_i) \\ \hat{\beta}_i^C(h_i) \end{bmatrix} \right\| \quad (13)$$

représentant notre pseudo-phase à une hauteur donnée. Notre choix d'appeler ces fonctions des *pseudo-phases* s'explique par le fait que  $\hat{c}_i^C(h_i)$  et  $\hat{s}_i^C(h_i)$  représentent des estimateurs pour  $\cos(\varphi_i(h_i))$  et  $\sin(\varphi_i(h_i))$ , si  $h_i$  correspond à la hauteur réelle du point physique.

En égalisant les pseudo-phases entre les caméras, on obtient pour  $r \geq 2$  caméras :

$$C_{MCPPr}(h_i) = \sum_{k=1}^r \left\| \begin{bmatrix} \hat{c}_i^k(h_i) \\ \hat{s}_i^k(h_i) \end{bmatrix} - \frac{1}{r} \sum_{l=1}^r \begin{bmatrix} \hat{c}_i^l(h_i) \\ \hat{s}_i^l(h_i) \end{bmatrix} \right\|_2^2. \quad (14)$$

Par similitude avec la PDP dans laquelle les motifs sont déphasés les uns par rapport aux autres de  $\frac{2\pi}{n}$ ,  $n$  étant le nombre de motifs projetés, la translation entre acquisitions est telle que le déphasage introduit est de l'ordre de grandeur de  $\frac{2\pi}{n}$  modulo  $2\pi$ .

La Figure 2 montre les courbes résultantes des coûts développées MCI4 et MCP4 pour quatre points de la grille de référence dans le Cdp.

## 4 Simulation

### 4.1 Données

Pour tester notre méthode, notre logiciel interne de simulation de notre équipement illustré à la Figure 1 nous a permis de synthétiser les images vues par les caméras d'une scène virtuelle éclairée par la projection d'un motif particulier. Le simulateur est basé sur le principe du *raycasting* pour modéliser le comportement des pixels des capteurs. Contrairement aux *raycasters* classiques utilisant des

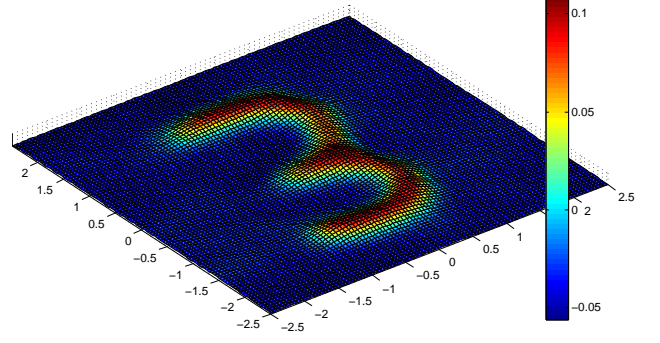


FIG. 4: Surface reconstruite par la méthode de Furukawa en utilisant les images de synthèse (idéales) avec projection d'un motif binaire aléatoire (visible dans Figure 3).

sources lumineuses simplifiées, celui-ci intègre la notion de vidéo-projecteurs. Les images sont alors calculées sur la base d'un modèle de réflexion lambertien.

De plus, le simulateur nous permet de modifier divers paramètres afin de pouvoir étudier la validité de notre approche dans le cas idéal et sa robustesse aux différentes sources d'incertitudes comme le bruit caméra et l'incertitude d'estimation du mouvement.

La scène virtuelle illustrée dans la Figure 3 est composée d'une surface planaire avec un "3" en relief et dont la texture représente un "D".

Pour chacune des caméras nous avons généré un lot de seize images sur lesquelles la scène a été translatée d'une image à l'autre en projetant un motif statique de franges sinusoidales. Nos méthodes sont alimentées par ces images. Pour situer nos méthodes par rapport à l'état de l'art nous les avons comparées à deux autres approches. La première des deux méthodes de référence est notre implémentation de l'approche PDP Statique (PDPS) de [10] pour laquelle nous avons synthétisé huit images par caméra avec une scène immobile mais des projections de franges déphasées. La seconde est l'implémentation de la méthode de stéréo passive de Furukawa et Ponce [2] pour laquelle nous avons généré seize images avec projections binaires aléatoires sur la scène en mouvement. Pour chaque déplacement nous obtenons une reconstruction éparse. Ces reconstructions sont ensuite fusionnées en remaillant l'union de leurs sommets. Pour ces trois cas nous ajoutons du bruit caméra ( $\sigma_{GL} \in \{0, 0.1275, \mathbf{0.255}, 0.51, 1.02\}$ ) sur une échelle de 0 à 255 niveaux de gris et du bruit sur le mouvement ( $\sigma_{ME} \in \{0, 0.0037, \mathbf{0.0075}, 0.015, 0.03\}$  millimètres); les niveaux de bruit de référence de notre système réel sont mis en gras. Dans les deux cas il s'agit de bruit gaussien additif.

### 4.2 Résultats

**Comparaison aux techniques à projection de motifs binaires aléatoires.** La Figure 4 montre une reconstruction basée sur les images à motifs aléatoires et reconstruites pas à pas avec le code de Furukawa et dont les multiples re-

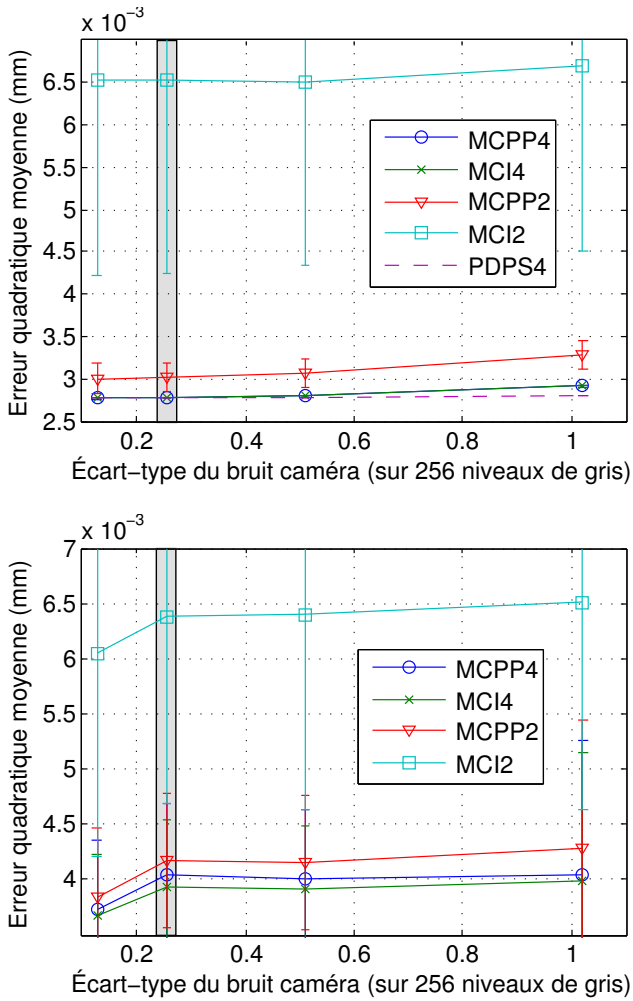


FIG. 5: EQM en fonction du bruit de caméra sans (haut) et avec (bas) incertitude de mouvement ( $\sigma_{ME} = 7.5\mu m$ ). Les mesures ont été moyennées sur 100 tirages.

constructions indépendantes ont été fusionnées.

Bien que le résultat soit visuellement très satisfaisant, celui-ci démontre que la méthode n'est pas adaptée à notre cas : l'essai dans le cas idéal (sans bruit caméra, ni d'incertitudes de mouvement) a montré une Erreur Quadratique Moyenne (EQM) de 0.0467mm des hauteurs par rapport à la vérité terrain et a demandé un paramétrage manuel pour trouver le meilleur compromis entre conservation de détails et suppression des points aberrants. Cette erreur dans le cas idéal est bien supérieure à celles de nos méthodes même sous les influences des différentes incertitudes (EQM maximale de 0.004mm). Par conséquent nous n'avons pas inclus cette méthode dans les évaluations qui suivent.

**Bruit caméra.** Les deux graphiques dans la Figure 5 montrent les EQM mesurées pour nos différentes fonctions de coûts et moyennées sur cent tirages aléatoires de bruit. Le graphique du haut montre l'EQM en fonction du bruit caméra avec estimation de mouvement idéale et inclut comme référence la scène reconstruite avec un PDPS

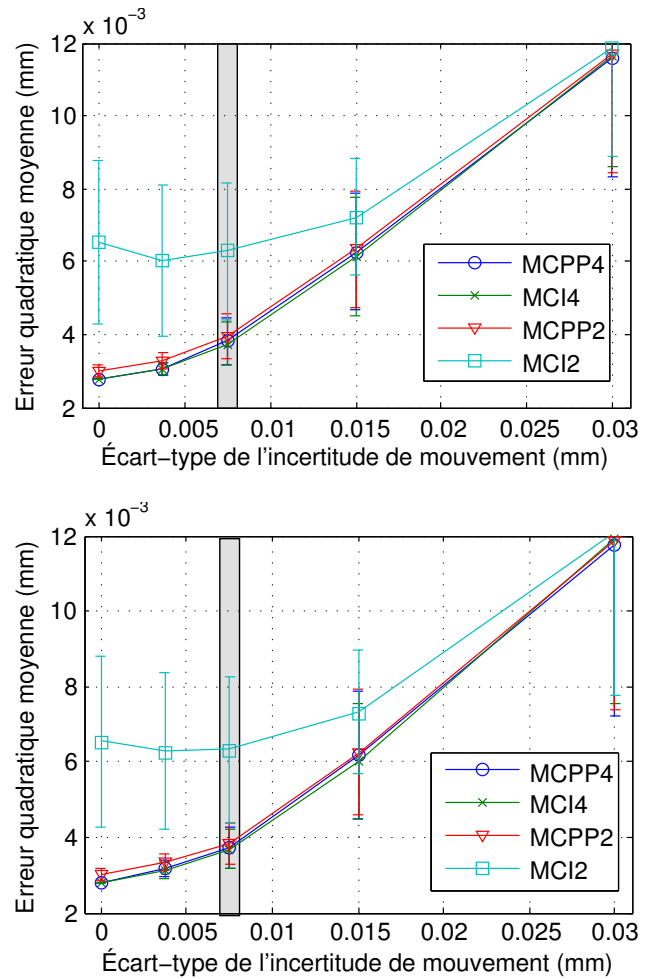


FIG. 6: EQM en fonction de l'incertitude de mouvement sans (haut) et avec (bas) bruit de caméra ( $\sigma_{GL} = 0.255$ ). Les mesures ont été moyennées sur 100 tirages.

utilisant quatre caméras (PDPS4) sur une scène statique. Nous voyons que MCPP4 et MCI4 sont légèrement plus sensibles au bruit caméra que PDPS4. MCPP2 est encore un peu plus sensible et MCI2 est bien moins robuste que les autres coûts.

Avec les incertitudes de mouvement de notre matériel (graphique du bas), les EQM de toutes nos méthodes augmentent légèrement et deviennent moins répétables (voir barres d'incertitude). L'EQM élevé même à faibles bruits peut s'expliquer par un nombre important de points aberrants provoqués par des minima locaux et/ou des faux minima.

La notion d'incertitude de mouvement n'a pas de sens pour PDPS4 et donc, ce dernier n'est plus inclus dans le deuxième graphique.

**Incertitudes de mouvement.** Cette expérience dévoile la sensibilité aux incertitudes d'estimation du mouvement des différents coûts. La Figure 6 montre l'EQM pour une incertitude de mouvement croissante sans (graphique du haut) et avec (graphique du bas) bruit caméra de référence.



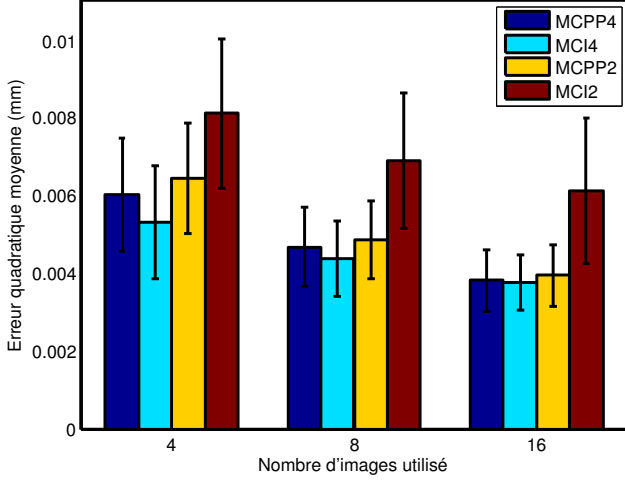


FIG. 7: EQM en fonction du nombre d'images d'entrée avec bruit de caméra et incertitude de mouvement de références, moyennée sur 100 tirages.

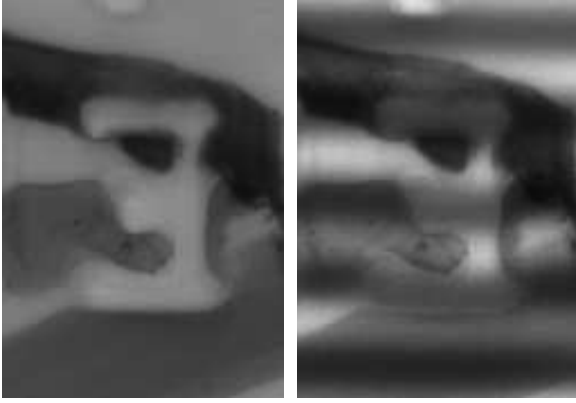


FIG. 8: Région d'intérêt d'une scène réelle avec éclairage uniforme (gauche) et projection de franges (droite).

Dans les deux cas les EQM moyennées de MCPP4 et MCI4 sont très proches l'une de l'autre et celles de MCPP2 est légèrement plus élevée. Les mesures pour MCI2 sont bien moins bonnes dans les deux cas dans cette expérience aussi. Par contre, les résultats de toutes les méthodes semblent converger pour des incertitudes de mouvement plus importantes.

**Nombre d'images.** Finalement nous étudions les variations des performances en fonction du nombre d'images utilisées sous l'influence du bruit caméra et de l'incertitude de mouvement de référence de notre matériel. La Figure 7 montre qu'en augmentant le nombre d'images, les coûts MCPP4, MCI4 et MCPP2 ont des performances de plus en plus équivalentes. MCI2 reste moins performant.

TAB. 1: Moyennes ( $\mu$ ) et écart-types ( $\sigma$ ) des différences des cartes de hauteurs par rapport à PDPS4.

	MCPP4	MCI4	MCPP2	MCI2
$\mu$ (mm)	-0.0052	-0.0144	-0.0040	-0.0127
$\sigma$ (mm)	0.0168	0.0352	0.0185	0.0342

## 5 Exemple Réel

### 5.1 Données

Une carte de crédit blanche sur laquelle nous avons rajouté de la texture éparsée en couleur avec des feutres masque l'application réelle mais conserve à peu près l'échelle, ainsi qu'une texture et une topologie ressemblante aux objets ciblés. La carte de crédit a été placée sur la cible réelle (dont la nature est confidentielle) et a passé un cycle d'acquisition de seize images pour chacune des quatre caméras et l'estimation du mouvement est donnée par le convoyeur robotique.

### 5.2 Résultats

La Figure 9 montre les reconstructions 3D obtenues. Les hauteurs sont encodées sur l'axe de couleur pour toutes les reconstructions à l'exception de la Figure 9b sur laquelle la texture reconstruite a été superposée.

La reconstruction de référence est donnée par la PDPS4 (Figures 9a et 9b) sans déplacement de la scène.

Les quatre reconstructions en mouvement (9c à 9f) sont cohérentes avec celle de la PDPS4. Surtout les reconstructions obtenues avec MCPP4 et MCPP2 y ressemblent au mieux. En effet, MCI $r$  est plus sensible à des changements de texture abruptes (e.g. en haut à gauche) que MCPP $r$ , ce qui cause des artefacts de reconstruction plus importants le long de ces contours. D'ailleurs, dans ce cas extrême, la reconstruction PDPS4 présente aussi des artefacts à cet endroit. En revanche, le contour au milieu à gauche provoque des artefacts importants dans MCI $r$ . Ils sont légers dans MCPP $r$  et inexistant dans la reconstruction PDPS4. Aux endroits à faible signal, par exemple dans les zones de couleur sombre (en haut à gauche), MCI $r$  souffre d'un bruit de reconstruction élevé, alors que MCPP $r$  reste relativement robuste. De plus, dans la région planaire en bas à droite, MCPP $r$  retourne une reconstruction aussi lisse que PDPS4 contrairement à MCI4.

La vallée (en haut à gauche) montre quelques artefacts dans la reconstruction obtenue par MCPP $r$  et de nombreux dans MCI $r$ . Nous supposons que cela est lié aux interrélaxions et/ou spéularités qui résultent du miroir convexe que représente la vallée.

La Figure 10 montre la distribution des différences de reconstruction entre nos méthodes et la PDPS4. Les modes des quatre reconstructions sont biaisés de  $-5\mu m$  à  $-10\mu m$ . Le Tableau 1 montre que les méthodes MCPP $r$  ressemblent plus à la référence (écart-types plus faibles) et sont moins biaisées (moyenne plus proche de 0) que les méthodes MCI $r$ .

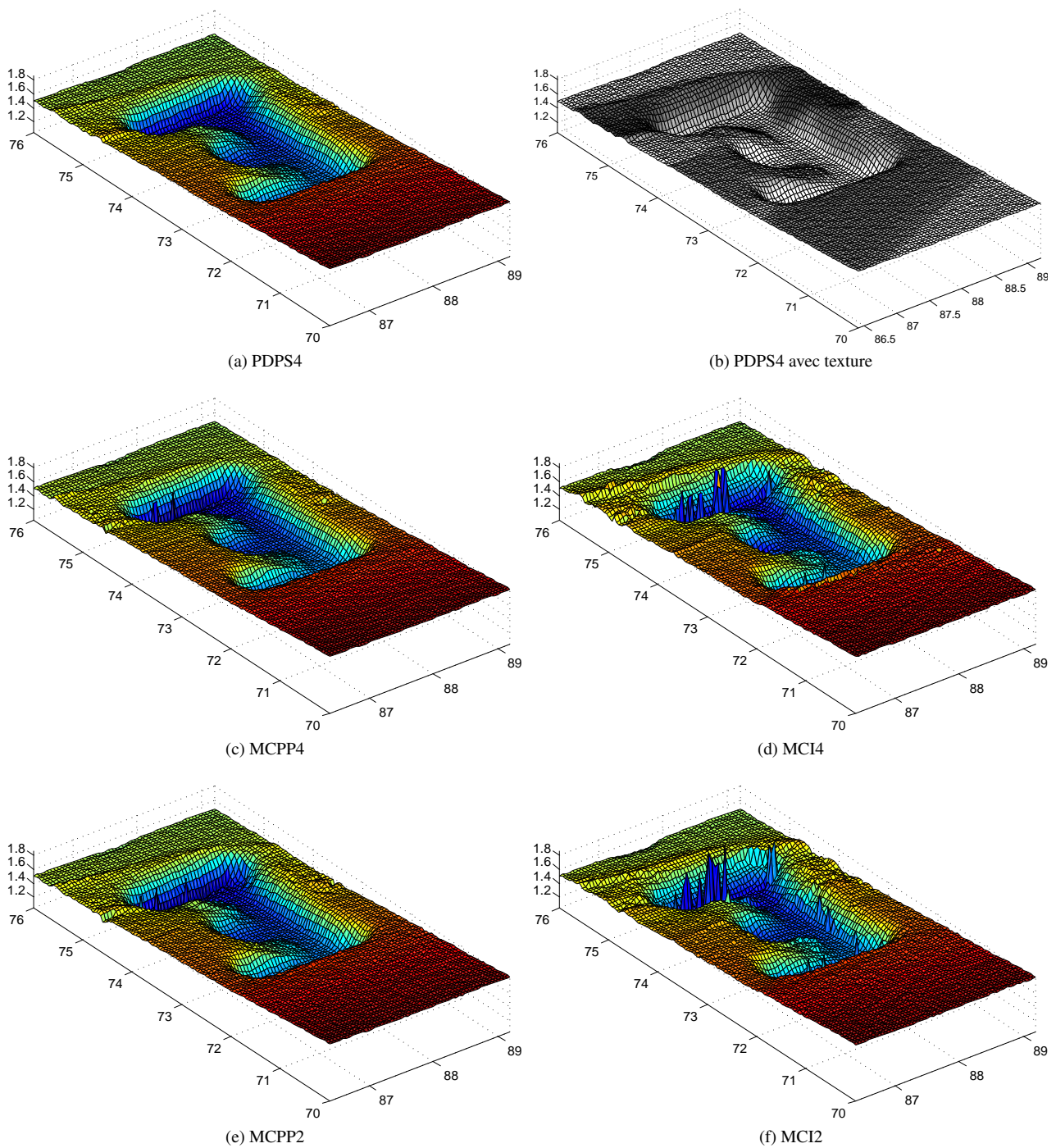


FIG. 9: Reconstructions de la scène réelle en utilisant toutes les quatre caméras ((a)-(d)) ou les caméras 1 et 3 ((e)-(f)).



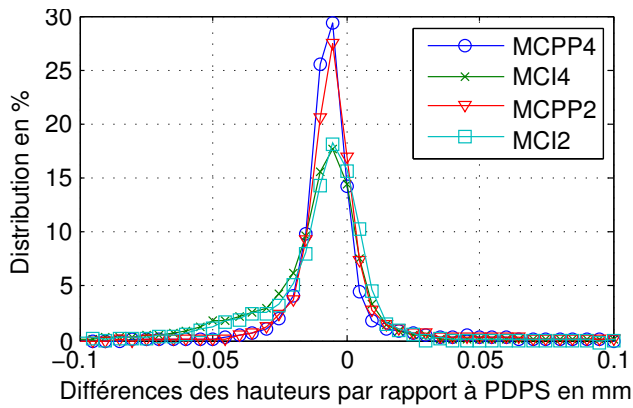


FIG. 10: Distributions des différences des cartes de hauteurs par rapport à PDPS4.

## 6 Discussion

Les approches introduites se distinguent par le fait qu'elles exploitent (MCPP $r$ ) ou non (MCI $r$ ) la connaissance *a priori* du motif projeté. MCI $r$  peut être considérée comme une approche spatio-temporelle avec illumination non structurée [13] avec un motif structuré. De ce fait nous considérons MCI $r$  comme une mise en oeuvre particulière de l'état de l'art. PPMx par contre représente une nouvelle réalisation dans la catégorie des méthodes de PDP en exploitant une connaissance des franges projetées et du déphasage induit par le mouvement de la scène.

Notre méthode est adaptée à un environnement industriel. Elle retourne des reconstructions denses de surfaces texturées en mouvement sans perte de résolution et sans exiger des caméras et vidéo-projecteurs hautes-fréquences. Elle permet même de réduire la complexité du système de projection car le motif à projeter ne varie pas pendant la séquence d'acquisition. Ceci limite les facteurs d'incertitudes du système comme la non-linéarité d'intensité projeté par le vidéo-projecteur ou les asymétries du motif sinusoïdal provoquées par des limitations de la technologie de projection.

De plus, notre système n'utilise pas d'optiques télécentriques qui coûtent bien plus chères que les optiques grand public que nous avons utilisé (optique d'appareil photo numérique d'entrée de gamme) et limitent considérablement le CdV ou font croître la taille du système optique.

## 7 Conclusion

Nous avons présenté une nouvelle méthode de reconstruction à projection de lumière structurée qui comble une lacune dans le domaine de la métrologie et de l'inspection industrielle entre les méthodes de PDP de l'état de l'art et des méthodes appelées "one-shot". Nous avons démontré la validité de la méthode avec des simulations synthétiques et nous avons donné un exemple sur des données réelles. De même, nous avons étudié les performances de notre solution qui se rapproche de celle de la PDPS et qui dépasse

celles des approches one-shot.

Dans de futurs travaux, nous envisageons d'approfondir les propriétés analytiques des fonctions de coût, d'exploiter la texture reconstruite [11], d'introduire une régularisation spatiale entre les points et d'étendre les fonctions de coût par un modèle de réflectance plus sophistiqué incluant les spécularités.

## Références

- [1] Davis, J., Nehab, D., Ramamoorthi, R., Rusinkiewicz, S. : Spacetime stereo : A unifying framework for depth from triangulation. PAMI 27(2), p. 296–302 (2005)
- [2] Furukawa, Y., Ponce, J. : Accurate, dense, and robust multi-view stereopsis. PAMI 32(8), p. 1362–1376 (2010)
- [3] Hall-Holt, O., Rusinkiewicz, S. : Stripe boundary codes for real-time structured-light range scanning of moving objects. ICCV (2001)
- [4] Huang, L., Kemao, Q., Pan, B., Asundi, A.K. : Comparison of Fourier transform, windowed Fourier transform, and wavelet transform methods for phase extraction from a single fringe pattern in fringe projection profilometry. Optics and Lasers in Eng. 48, p. 141–148 (2010)
- [5] Jongenelen, A., Bailey, D., Payne, A., Dorrington, A., Carnegie, D. : Analysis of errors in tof range imaging with dual-frequency modulation. Instrumentation and Measurement 60(5), p. 1861–1868 (2011)
- [6] Konolige, K. : Projected texture stereo. ICRA, p. 148–155 (2010)
- [7] Salvi, J., Fernandez, S., Pribranic, T., Llado, X. : A state of the art in structured light patterns for surface profilometry. Pattern Recognition 43(8), p. 2666–2680 (2010)
- [8] Smisek, J., Jancosek, M., Pajdla, T. : 3d with kinect. ICCV Works., p. 1154–1160 (2011)
- [9] Taguchi, Y., Wilburn, B., Zitnick, C. : Stereo reconstruction with mixed pixels using adaptive over-segmentation. CVPR (2008)
- [10] Wang, Y., Liu, K., Hao, Q., Lau, D., Hassebrook, L.G. : Multicamera phase measuring profilometry for accurate depth measurement. SPIE, vol. 6555 (2007)
- [11] Wang, Y., Liu, K., Hao, Q., Wang, X., Lau, D., Hassebrook, L. : Robust active stereo vision using kullback-leibler divergence. PAMI 34(3), p. 548–563 (2012)
- [12] Weise, T., Leibe, B., Gool, L.V. : Fast 3d scanning with automatic motion compensation. CVPR (2007)
- [13] Zhang, L., Curless, B., Seitz, S. : Spacetime stereo : shape recovery for dynamic scenes. CVPR 2, p. 367–74 (2003)
- [14] Zhang, L., Curless, B., Seitz, S.M. : Rapid shape acquisition using color structured light and multi-pass dynamic programming. 3D Data Proc. Visual. and Trans. (2002)
- [15] Zhang, Z. : A flexible new technique for camera calibration. PAMI 11(22), p. 1330–1334 (2000)